

# Knowledge Representation in Conceptual Realism

Nino B. Cocchiarella  
Department of Philosophy  
Indiana University  
Bloomington, IN 47405  
U.S.A.

Knowledge representation in A.I. involves more than the representation of a large number of facts or beliefs regarding a given domain, i.e. more than a mere listing of those facts or beliefs as data structures. It may involve, for example, an account of the way the properties and relations that are known or believed to hold of the objects in that domain are organized into a theoretical whole—such as the way different branches of mathematics, or of physics and chemistry, or of biology and psychology, etc., are organized, and even the way different parts of our commonsense knowledge or beliefs about the world can be organized.<sup>1</sup> But different theoretical accounts will apply to different domains, and one of the questions that arises here is whether or not there are categorial principles of representation and organization that apply across all domains regardless of the specific nature of the objects in those domains. If there are such principles, then they can serve as a basis for a general framework of knowledge representation independently of its application to particular domains. In what follows I will give a brief outline of some of the categorial structures of conceptual realism as a formal ontology. It is this system that I propose we adopt as the basis of a categorial framework for knowledge representation.

---

<sup>1</sup>See, e.g., Hobbs and Moore 1985.

# 1 Predicate Logic

Predicate logic, by which I mean higher-order and not just first-order predicate logic, provides the clearest and best example of a categorial structure for organizing the way objects are represented in terms of the properties and relations that hold between them. Predication is after all the fundamental nexus of both language and thought—and, at least from the ontological perspective of conceptual realism, of reality as well.

In first-order predicate logic there is a categorial distinction between predicate expressions (which represent verb phrases of natural language), all of which have a certain (finite) number of subject or argument positions as part of their functional role as predicates, and singular terms, e.g. individual variables and other expressions that can occupy the subject or argument positions of predicate expressions. But the quantifiers of first-order logic reach only into the subject or argument positions that are part of the functional role of predicate expressions, and not into the positions of predicates as well. The variables to which these quantifiers are affixed are called individual variables, which, because they have all and only the objects in the domain of discourse as their values, are understood in formal ontology to represent the basic ontological category of objects. For that reason, the quantifier phrases generated with these variables are called *objectual quantifiers*.

In second-order predicate logic there are also predicate variables that have predicate expressions (complex or simple) as their substituends, and quantifiers are allowed to bind those variables and reach into predicate positions as well. In nominalism, as a theory of predication (or formal ontology), only first-order, objectual quantifiers have ontological significance, which means that predicate quantifiers (and variables) are either omitted altogether or given only a substitutional interpretation. The latter option brings with it constraints on the logic of predicate quantifiers that in effect exclude *impredicative* formulas, i.e. formulas containing bound predicate variables, from being substituends of predicate variables (i.e. substituends that preserve validity according to nominalism). The system that results by imposing these constraints is known as (standard) *predicative* second-order logic. (Nominalism can be extended into standard ramified predicative second-order logic in which formulas with bound predicate variables of a given level of ramification can be substituends of predicates of any higher level; but it still excludes a comprehension principle in which formulas with bound predicate variables of a given level can be substituends of the predicate variables of that or any

lower level.<sup>2</sup> )

Nominalism is too restrictive a framework for knowledge representation in that it fails to take into account both the conceptual nature of representation and the ontological structure of the world in terms of which our knowledge and beliefs about a given domain are organized. Conceptualism goes beyond nominalism in its recognition of predicable concepts as values of predicate variables (with respect to which predicate quantifiers have referential significance), and conceptual realism goes even further in its recognition of the contents of our concepts as intensional objects (denoted by nominalized predicates) and of the possibility of there being properties, relations, and natural kinds in nature corresponding to some, even if not all, of our concepts.

Concepts in the sense intended here are intersubjectively realizable (and in that sense *objective*) cognitive structures that underlie our ability to think and communicate with one another. They are not objects (and therefore cannot be values of individual variables) but are rather unsaturated cognitive structures that are realized (saturated) in particular mental acts – including speech acts as overt forms of mental acts having a communicative role. Predicable concepts in particular are based on cognitive capacities to identify, characterize and relate objects in various ways, and they underlie our ability to follow the rules of language regarding the use of predicate expressions. It is through the exercise of these capacities that our mental acts (and therefore our speech acts as well) are informed with a predicable nature.

There are two major types of conceptualism, depending on the constraints for concept-formation. The first is a constructive conceptualism, which is so-called because it imposes a constraint that precludes *impredicative concept-formation*, i.e. the formation of concepts on the basis of a totality to which those concepts belong or form a part. The logic resulting from these constraints is a nonstandard form of predicative second-order logic (which can be extended into a nonstandard form of ramified second-order logic), which is similar and yet different from the standard predicative second-order logic of nominalism. The second is a holistic conceptualism in which impredicative concept-formation is allowed, but without nullifying the (ramified) predicate quantifiers that range over the so-called *predicative* concepts of constructive conceptualism. It is only this type of conceptualism that pro-

---

<sup>2</sup>See Church 1956, section 58, for a formal description of standard predicative second-order logic, and Cocchiarella 1986a, chapter one, for a semantical formulation of nominalism in terms of such a logic.

vides an adequate framework for knowledge representation of those parts of mathematics that essentially involves impredicative concepts—such as the concepts involved in the continuity schema or least upper bound principle of real number theory. It is only holistic conceptualism that we have in mind when speaking of conceptualism hereafter.

The universal and existential quantifiers,  $\forall$  and  $\exists$ , when affixed to predicate variables, will be understood, accordingly, to refer to predicable concepts in general, whether predicative or impredicative. In those domains of knowledge representation (such as various branches of mathematics) where the predicative-impredicative distinction is important, we can use the quantifiers  $\forall!$  and  $\exists!$  to refer, when affixed to predicate variables, exclusively to the predicative/constructive concepts of the domain in question (as determined by the logic of constructive conceptualism). In domains where this distinction is not important, we can simply ignore or omit all formulas containing  $\forall!$  or  $\exists!$  as constituents.<sup>3</sup>

In domains in which the distinction is important, incidentally, we can also add an epistemic operator  $\Box^k$  (and suitable axioms) to represent our constructive knowledge regarding the objects in those domains. This operator—read as ‘it is constructively knowable that’—can be used to express various epistemological principles (much in the manner in which Gödel used the  $\Box$  of S4 modal logic to interpret sentential intuitionist logic), such as the thesis that objects will fall under a constructive concept only if it is constructively knowable that they do. That is, formally,

$$(K!) \quad (\forall!F^n)(\forall x_1)\dots(\forall x_n)[F(x_1, \dots, x_n) \rightarrow \Box^k F(x_1, \dots, x_n)],$$

can be taken as a valid epistemological principle of conceptualism that distinguishes the role of constructive/predicative concepts from predicable concepts in general. The same thesis with  $(\forall!F^n)$  replaced by  $(\forall F^n)$  is not valid because the latter quantifier refers to nonconstructive/impredicative concepts as well. (In general, we use capital Roman letters  $F, G, H$  and  $R$ , with or without numerical subscripts, as predicate letters, and  $x, y, z$ , with or without subscripts, as individual variables.) Accordingly, because the complement of a predicative concept is also a predicative concept, the following *Principle of Rigidity* for predicative concepts,

$$(PR!) \quad (\forall!F^n)(\forall x_1)\dots(\forall x_n)[\Box^k F(x_1, \dots, x_n) \vee \Box^k \neg F(x_1, \dots, x_n)],$$

---

<sup>3</sup>See Cocchiarella 1986b for the formal details of the logic of constructive and holistic conceptualism.

is valid in the general domain of constructive knowledge (as understood in constructive conceptualism).

The question as to which primitive predicate constants stand for constructive/predicative concepts in a given domain can be decided only by external criteria. For example, if  $P$  is a predicate constant that is to stand for a predicative concept of a given domain as decided by some external criteria, we can express this fact internally by means of a meaning postulate of the following form:

$$(\exists! F^n) \Box^k (\forall x_1) \dots (\forall x_n) [F(x_1, \dots, x_n) \leftrightarrow P(x_1, \dots, x_n)].$$

Which complex formulas represent predicative concepts of the domain is then determined by the comprehension principle and other laws of the predicative logic of constructive conceptualism.

## 2 Conceptual Intensional Realism and the Logic of Nominalized Predicates

Higher-order predicate logic goes beyond the impredicative second-order logic of holistic conceptualism by introducing a syntactical operation that transforms predicates (and formulas) into abstract singular terms, i.e. into nominalized predicates (and formulas). The operation is indicated here by the deletion of the parentheses (and commas in the case of a relational predicate) that are part of a predicate expression in its functional role as a predicate. For example, in addition to the role of  $F$  in  $F(x)$  as a predicate expression, we also have the role of the nominalization of  $F$  as an abstract singular term in  $G(F)$  and  $R(y, F)$ —as well as in  $F(F)$ , where  $F$  occurs first with a pair of parentheses as a predicate expression, and then without the parentheses as a singular term. For complex predicates, as determined by formulas, we will use the  $\lambda$ -operator (which binds the variables affixed to it.) Thus, where  $\phi$  is a formula, and  $x_1, \dots, x_n$  are distinct individual variables,  $[\lambda x_1 \dots x_n \phi]$  is an  $n$ -place predicate expression, which in its role as a predicate brings with it a pair of parentheses (and commas if  $n > 1$ )—as, e.g. in its occurrence in  $[\lambda x_1 \dots x_n \phi](y_1, \dots, y_n)$ . The nominalization of such a complex predicate expression is represented the same way as for simple predicate letters; namely, by deleting the parentheses (and commas) that are a functional part of its role as a predicate – as, e.g., in  $R(y, [\lambda x \phi])$ , as well as in  $[\lambda x \phi]([\lambda x \phi])$ , where

the one-place predicate expression  $[\lambda x\phi]$  occurs both as a predicate and as an abstract singular terms. (By means of  $\lambda$ -abstraction, a formula  $\phi$  can be construed as a 0-place predicate,  $[\lambda\phi]$ , which, as an abstract singular term, is read as ‘that  $\phi$ ’. For convenience, we write ‘ $[\phi]$ ’ instead of ‘ $[\lambda\phi]$ ’. It is in this way that we will represent nominalized formulas hereafter.)

What a nominalized predicate denotes, if it denotes anything at all, cannot be the concept that the predicate stands for in its role as a predicate—because concepts, as we have said, are unsaturated cognitive structures, and as such they are not objects. In conceptual realism what a nominalized predicate denotes is an intensional object—namely, the intension or content of the concept that the predicate stands for in its role as a predicate. Here, by the intension or content of a concept I mean a hypostatization, reification, or projection into the domain of objects of the truth conditions determined by the different possible applications of that concept. It is by means of such a projection, or conceptual nominalization, that we purport to denote the intension or content of a concept as if it were an independently existing real Platonic form. Thus, for example, not only do we predicate of someone that she is kind and wise, or of a box that it is green and rectangular, but also, through conceptual nominalization of the concepts that we predicate, we purport to denote the properties (in the Platonic sense) of *being kind* and *being wise*, or the properties of *being green* and *being rectangular*—i.e. the properties *kindness*, *wisdom*, *greenness*, and *rectangularity*. It is through conceptual nominalization (as a product of cultural evolution) that we hypostatize or reify a concept and are able to grasp its content or intension as an (abstract) object by starting out from the concept as a cognitive capacity. The assumption that conceptual nominalization (reification) leads to real abstract objects (as emergent products of cultural evolution) is an ontological posit that goes beyond conceptualism proper, and it forms part of what I mean by conceptual realism as a formal ontology. This part of conceptual realism is called conceptual intensional realism.<sup>4</sup>

The addition into impredicative second-order logic of nominalized predicates as abstract singular terms was the move that Frege made in going from the system of his *Begriffsschrift* to that of his *Grundgesetze*. (He did this, in effect, through his use of the spiritus lenis (or smooth breathing  $\acute{e}$ ),

---

<sup>4</sup>In Brachman 1985 it is claimed that the knowledge representation programming language KL-ONE needs a “metadescription” level in which to deal with KL-ONE concepts as “internal entities” (198). Such a “level” is contained in the logic of nominalized predicates that is part of the system of conceptual intensional realism.

as an abstraction operator. Thus, instead of  $F(F)$ , or  $F([\lambda x F(x)])$ , as in our notation, Frege used  $F(\dot{\epsilon}F(\dot{\epsilon}))$ .) This addition was important and fundamental to Frege's development of arithmetic as part of logic, by which he meant only extensional logic. That is why the objects he took nominalized predicates to denote in his system were classes (*Begriffsumfänge*) or value-ranges (*Wertverläufe*), rather than the intensional objects that we take them to be in conceptual realism. Our intensional approach does not affect Frege's reduction of arithmetic to logic, and in fact it opens a new and important way of developing knowledge representation in intensional mathematics.<sup>5</sup> Indeed, in conceptual realism (as I understand it here) all abstract objects are assumed to be concept-correlates (i.e. hypostatized, or reified, contents of concepts) that have their being in the concepts whose correlates they are. (But this does not mean that all existential posits regarding such objects can be justified on purely conceptual or logical grounds.)

A logic of nominalized predicates is important not only for an analysis of mathematical objects but also for an adequate representation of the nominalized predicate expressions of natural language and of the things we say in terms of such expressions. It is by means of such a logic, as I have explained elsewhere, that a formal account can be given of the intensional verbs of natural language, including in particular the verbs that represent propositional attitudes.<sup>6</sup> It is also in this way that a logical analysis of fiction, and of the content of stories in general, whether true or false, can be given. A comprehensive system of knowledge representation should include not only the conceptual structures of (impredicative) second-order predicate logic, accordingly, but also the logic and ontology of nominalized predicates (and sentences) as well.

The one problem with this proposal is Russell's paradox (which is a special case of Cantor's paradox). That is, with nominalized predicates and an impredicative comprehension principle, Russell's paradox is derivable from the following instance of the comprehension principle:

$$(\exists F)(\forall x)[F(x) \leftrightarrow (\exists G)(x = G \wedge \neg G(x))].$$

Russell's solution to his paradox was the theory of ramified types, which, for most purposes, could be simplified to the theory of simple types. It is the latter system that is now generally referred to as higher-order predicate

---

<sup>5</sup>See Shapiro 1985 for a survey of some approaches to intensional mathematics.

<sup>6</sup>See Cocchiarella 1989b, sections 7-9.

logic. It was Church’s version of the theory of simple types, for example, that Montague developed into an intensional logic that became the semantical framework for natural language in what is now known as Montague grammar.<sup>7</sup>

There are a number of problems with the theory of types as a framework for knowledge representation, not the least of which is its inability to provide an adequate account of certain meaningful sentences of natural language that violate the restrictions of type theory. Fortunately, these problems can be overcome by shifting type restrictions from the object-language to the metalanguage—in particular to the conditions of well-formedness for  $\lambda$ -abstracts as complex predicates. The result of imposing these restrictions on the  $\lambda$ -abstracts of second-order predicate logic with nominalized predicates as abstract singular terms is known as the system  $\lambda HST^*$ , which I have shown to be equiconsistent with the theory of simple types.<sup>8</sup> G. Chierchia has subsequently shown that  $\lambda HST^*$  (supplemented with tense and modal operators) can be used in Montague grammar in place of Montague’s type-theoretical intensional logic.<sup>9</sup>

The restriction on  $\lambda$ -abstracts does not mean that formulas of the form  $F(F)$ , or, as in Russell’s argument,  $\neg F(F)$ , are meaningless in  $\lambda HST^*$ . In fact,  $(\exists F)F(F)$  is provable in  $\lambda HST^*$  (for  $F = [\lambda x(x = x)]$ ). Nevertheless, it is inappropriate in conceptualism that the complex predicate for the Russell concept—namely,  $[\lambda x (\exists G)(x = G \wedge \neg G(x))]$ —should not be well-formed. It is, after all, through the construction of this concept that we are able even to think of Russell’s argument. Fortunately, the restriction does not apply to another system I have formulated, namely,  $HST_\lambda^*$ , that is closely related to  $\lambda HST^*$ . In  $HST_\lambda^*$ , the  $\lambda$ -abstract for the Russell concept is well-formed, as are all  $\lambda$ -abstracts of second-order logic with nominalized predicates. Russell’s paradox is avoided in this system not by rejecting the Russell concept but by denying that the nominalization of such a concept must denote. That is, although the Russell concept “exists” qua concept, i.e.

$$(\exists F) ([\lambda x (\exists G)(x = G \wedge \neg G(x))] = F)$$

is provable in  $HST_\lambda^*$ , nevertheless, because of the truth conditions it determines, this concept cannot be nominalized (i.e. hypostatized/reified) so as

---

<sup>7</sup>See Montague 1974, chapters 7 and 8.

<sup>8</sup>See Cocchiarella 1986a.

<sup>9</sup>See Chierchia 1984 and Chierchia 1985.



to result in an *object*, i.e.,

$$\neg(\exists y) ([\lambda x (\exists G)(x = G \wedge \neg G(x))]) = y)$$

is also provable in  $HST_\lambda^*$ . (This means that, unlike  $\lambda HST^*$ , the logic of  $HST_\lambda^*$  is free of existential presuppositions regarding singular terms. Conceptual nominalization only purports to denote the result of a projection or hypostatization of the truth conditions of a concept – and such a projection or hypostatization cannot succeed if those truth conditions lead to a contradiction.) The two systems  $\lambda HST^*$  and  $HST_\lambda^*$ , as I have shown elsewhere (in 1986a), are equiconsistent, which means that  $HST_\lambda^*$  is also equiconsistent with the theory of simple types, and that therefore it is as good as  $\lambda HST^*$  as the basis of an intensional logic for Montague grammar. It is  $HST_\lambda^*$ , supplemented (when needed) with the predicative logic of constructive conceptualism, that we propose to take as the basic predicate logic for knowledge representation.

### 3 Tense and Modal Logic

A fundamental feature of all conceptual (and A.I.) activity *as a temporal process* are the internal cognitive schemata by which we orientate ourselves in time. Such schemata are implicit in the form and content of all our mental acts. In mathematical or strictly scientific theories or domains of knowledge representation, these schemata are usually replaced by a tenseless (‘at time  $t$ ’) idiom based upon some externally imposed metric. Such an idiom is not appropriate for the representation of any form of conceptual activity in which it is necessary to distinguish between the past, the present, and the future as internal features of our mental acts, including, in particular, natural language understanding. Indeed, in natural language with its dating and temporal language of years, months, days, minutes, etc., the distinction between the past, the present and the future remains an essential component of the tenses of our verbs. We say, for example, that Columbus *discovered* America in 1492, and that the republicans *will* try to regain the U.S. presidency in 1996, etc., so that even the inclusion of a date (and time) does not affect our intrinsic temporal orientation in terms of the past, the present and the future.

One way to represent the internal cognitive schemata by which we orient ourselves in time is by means of the tense operators  $\mathcal{P}$  and  $\mathcal{F}$ , which are read

‘it was the case that’ and ‘it will be the case that’, respectively. There are other tense operators we might introduce as well, such as  $\mathcal{N}$ , for ‘it is now the case that’, or  $\mathcal{S}$  and  $\mathcal{U}$  for ‘since’ and ‘until’; but we will restrict ourselves here to just  $\mathcal{P}$  and  $\mathcal{F}$ .<sup>10</sup> The simple present tense, which (given quantification over past and future objects) covers most of the purposes of  $\mathcal{N}$ , is implicit in all basic form of predication. The operators  $\mathcal{P}$  and  $\mathcal{F}$  are then understood to modify the simple present tense into the past and future tense, respectively; and iterated applications of these operators are then understood to represent the past of the past, the future of the past, the past of the future, the future of the future, etc.

Different assumptions about the structure of time, e.g. that it is dense, discrete, continuous, that it has a beginning, an end, or neither, etc., can be represented in terms of these operators. We will assume as invariant for knowledge representation only the tense-logical truths of a serially ordered local time and leave additional assumptions about the structure of time to special domains of knowledge representation.<sup>11</sup> For purposes of special relativity theory, other local times can be brought into our analyses through the introduction of causal tense operators  $\mathcal{P}_c$  and  $\mathcal{F}_c$ , which can be read as ‘it causally was the case that’ and ‘it causally will be the case that’, and which are based upon a (light) signal relation between the momentary states of the continuants underlying different local times. Unlike the logical structure of the standard tenses, which represent a serial ordering, the logical structure of these tenses represent only a (strict) partial ordering. It is in terms of these causal tense operators, and not the standard tense operators (representing the linearity of a local time), that we can represent the possibility of special relativity theory in which a state of affairs can come (in the causal future of the posterior cone) to have been the case (in the causal past of the then prior cone) without its ever actually being the case. That is,  $(\mathcal{F}_c\mathcal{P}_c\phi \wedge \neg[\mathcal{P}_c\phi \vee \phi \vee \mathcal{F}_c\phi])$  is consistent in special relativity theory, even though  $(\mathcal{F}\mathcal{P}\phi \wedge \neg[\mathcal{P}\phi \vee \phi \vee \mathcal{F}\phi])$  is not.<sup>12</sup>

Tense logic also provides the means by which we can represent various alethic modal notions. (Cognitive modalities, such as belief or desire, are analyzed in terms of intensional objects, and not states of affairs the way

---

<sup>10</sup>See Prior 1967 for a general account of tense logic.

<sup>11</sup>See Cocchiarella 1991 for a complete axiom set for tense-logical truth as understood here.

<sup>12</sup>See Cocchiarella 1984, section 15, for a set-theoretic semantics of the difference between the standard and the causal tense operators.

alethic modalities are.) Indeed, some of the first alethic modal notions ever described in the history of thought are based on the cognitive schemata characterizing our orientation in time. The Megarian logician, Diodorus, for example, argued that what is possible is what either is or will be the case, and that what is necessary is what is and always will be the case<sup>13</sup> :

$$\begin{aligned}\Diamond^f \phi &=_{df} \phi \vee \mathcal{F} \phi, \\ \Box^f \phi &=_{df} \phi \wedge \neg \mathcal{F} \neg \phi.\end{aligned}$$

For Aristotle, the possible included the past as well as the present and the future; i.e. what is possible, according to Aristotle, is what either was, is, or will be the case (and therefore what is necessary is what always was, is, and always will be the case):

$$\begin{aligned}\Diamond^t \phi &=_{df} \mathcal{P} \phi \vee \phi \vee \mathcal{F} \phi, \\ \Box^t \phi &=_{df} \neg \mathcal{P} \neg \phi \wedge \phi \wedge \neg \mathcal{F} \neg \phi,\end{aligned}$$

Similar modal notions can be defined in terms of the causal tense operators, and a variety of other modal notions as well (such as a temporal notion of necessity based on the way the past has been).

In addition to the alethic modalities that can be analyzed in terms of tense operators, we also need a modality that goes beyond the actual events and states of affairs that make up the space-time manifold of our universe—but which does not also go beyond the causal structure of the world as determined by the laws of nature. That is, we need a causal or natural necessity to represent what is invariant across all possible worlds that have the same laws of nature as ours, and which in that regard is part of the *natural realism* that is implicit in our scientific theories, and explicit in our commonsense understanding of the world. The logic of natural realism is another component of what I mean by conceptual realism as a formal ontology, and for the purposes of developing this component, I will use  $\Box^c$ , and its dual  $\Diamond^c$ , as modal operators for causal or natural necessity and its dual possibility.

---

<sup>13</sup>See Prior, *op. cit.* for a discussion of Diodorus's argument.

## 4 The Logic of Existence as Part of the Logic of Being

In addition to the different temporal modalities that are implicit in the intrinsic temporal form of our conceptual activity, there are also different modes or kinds of being in the general category of objects. There are, for example, past objects, i.e. objects that did exist but no longer exist, and future objects that have yet to exist. That is, *there are* objects that do not (at present) exist, but which either *did exist* or *will exist*, and in that regard such objects have *being* even if they do not (now) also have *existence*.

There are nonexistent objects of other modes of being as well that we should be able to distinguish in special domains of knowledge representation. There are, for example, objects that are merely possible in nature (such as the oak tree that an acorn could have grown into before that acorn was crushed underfoot)—i.e. objects that could have existed in nature, but in fact never will exist at all (in our space-time universe). And then there are also abstract objects, i.e. objects that because of their abstract nature *cannot* exist (in the space-time manifold of any causally possible world), but which nevertheless have being as objects (i.e. as values of bound individual variables) – which is why we say *there are* such objects. Indeed, we can even say in indirect forms of speech that such objects “exist”—as when we say that someone believes that Pegasus exists, or when we say of fictional characters that they “exist” in the abstract world of a certain story. But this sort of abstract “existence” is not the same as the existence in a concrete possible world – because it is only as constituents of propositions (such as the propositions that make up the content of someone’s beliefs, or the content of a story, etc.) that abstract objects can be said to exist (in an indirect form of speech). Here, by a *proposition* we mean the type of intensional object that is denoted by a nominalized sentence (as opposed to the state of affairs that might correspond to the sentence itself). Propositions in this sense make up the abstract worlds of our theories, stories, beliefs, etc., which, unlike the concrete worlds made up of events and states of affairs, are usually incomplete, and in some cases even inconsistent. Existence only in an abstract world made up of propositions is not real existence, which is why we say that although abstract objects have being as objects they nevertheless do not, indeed cannot, exist.

A general framework for knowledge representation should be able to make all of these distinctions. In conceptual realism, we use the one-place predi-

cate constant  $E!$  to stand for the concept of (concrete) existence—so that ‘ $x$  exists’ (in the sense of concrete actuality) can be symbolized as ‘ $E!(x)$ ’. The predicates for past and future existence can be represented, accordingly, by  $[\lambda x \mathcal{P}E!(x)]$  and  $[\lambda x \mathcal{F}E!(x)]$ , and existence at some time or other in the history of the universe can be represented by  $[\lambda x \Diamond^t E!(x)]$ . Abstract existence (within the category of objects) can be defined negatively as the being of an object that cannot exist (as a concrete actuality), i.e. as  $[\lambda x \neg \Diamond^e E!(x)]$ . Restricted quantifiers for existents can be defined in terms of objectual quantifiers and  $E!$  as follows:

$$(\forall^e x)\phi \quad =_{df} \quad (\forall x)[E!(x) \rightarrow \phi],$$

$$(\exists^e x)\phi \quad =_{df} \quad (\exists x)[E!(x) \wedge \phi].$$

The distinction between being and existence (within the category of objects) is present in natural language not only in the distinction between the quantifier phrases ‘there *are* (*is*, *be*)’ and ‘there *exist*(*s*)’, but also in the distinction between those verb phrases (predicates) that entail existence (i.e. that can be true only of objects that exist) and those that do not. An object cannot be green, or red, or blue, etc., for example, unless it has a spatio-temporal location and exists as a concrete object. Similarly an object cannot be a horse (winged or not), or a mountain (made entirely of gold or not), unless it exists as a concrete object. But an object can be believed to be a (winged) horse, or it can be believed to be a (gold) mountain, and still not exist (in the space-time manifold), in which case its mode of being is that of an intensional object of someone’s belief space. That is, unlike being a (winged) horse, or being a (gold) mountain, *being believed to be a (winged) horse*, or *being believed to be a (gold) mountain*, does not entail existence. Similarly, *being a number* also does not entail (concrete) existence, but that does not mean that *there are* no numbers, i.e. that numbers do not have being as objects. And then too, *being an ancestor of Bill Clinton*, or *being a future female president of the U.S.A.*, also do not entail existence, in that many of Clinton’s ancestors no longer exist, and none of the future female presidents of the USA may yet have been born, which indicates that it is not only the verb phrases (predicates) for abstract objects that do not entail existence, but also phrases for past and future objects as well.

The distinction between verb phrases (predicates) that do, and verb phrases that do not, entail existence is determined by the predicable concepts that such phrases stand for – because it is the concepts that such

phrases stand for that determines their truth conditions. In one form or another, this distinction has been noted by medieval and nineteenth century logicians, such as, e.g., Bolzano.<sup>14</sup> It is implicit in Meinong's theory of objects, where it corresponds to his distinction between nuclear (*konstitutorisch*) and extra-nuclear (*ausserkonstitutorisch*) properties.<sup>15</sup> Formally, in conceptual realism, where concepts that entail existence are called *e-concepts*, we use the restricted quantifiers  $\forall^e$  and  $\exists^e$  to refer to e-concepts when affixed to predicate variables. That is, unlike the predicate quantifiers  $(\forall F^n)$  and  $(\exists F^n)$ , which range over all ( $n$ -ary) concepts,  $(\forall^e F^n)$  and  $(\exists^e F^n)$  range only over ( $n$ -ary) e-concepts.

Existence itself, it should be noted, is itself an e-concept, but one that is quite different from most ordinary e-concepts (such as being green, or being a horse, etc.), which involve being in the world in a specific concrete way. This is because, unlike most ordinary e-concepts, existence is an impredicative e-concept in that to exist is the same as to fall under an e-concept (of the ordinary sort). Instead of being taken primitively, accordingly, existence can be defined impredicatively (in terms of a totality to which it belongs) as follows:

$$E!(x) \quad =_{df} \quad (\exists^e F)F(x).$$

The logic of existence, as developed in terms of the quantifiers  $\forall^e$  and  $\exists^e$  affixed to both predicate and individual variables, can be developed either as a form of actualism or as part of possibilism, which can in turn be part of a more general framework such we have in mind here.<sup>16</sup> In those domains of knowledge representation where the distinction between e-concepts and concepts in general is not important, the second-order part of this logic can be ignored or omitted. But in some domains, such as the semantics and ontology of fiction in Parsons's reconstruction of Meinong's theory of objects, the distinction is important in the way it shows how Meinongian objects, and fictional objects in particular, can be interpreted in terms of properties of e-properties as intensional objects, i.e. in terms of the intensions of e-concepts.<sup>17</sup>

---

<sup>14</sup>See Bolzano 1972 (first published in German in 1837), p.xxxi.

<sup>15</sup>See Cocchiarella 1987, chapter 3, for a reduction of Meinong's nuclear-extranuclear distinction (as reconstructed in Parsons 1980) in terms of the difference (in logical realism) between e-properties and properties in general.

<sup>16</sup>See Cocchiarella 1991 and 1989a, sections 10-13, for a description of this logic in terms of actualism and possibilism.

<sup>17</sup>See Parsons 1980 and Cocchiarella 1987, chapter 3. Also, see Landini 1990 for dis-

## 5 Conceptual Natural Realism

As intensional objects, properties and relations are really not universals in the sense of predicable entities, i.e. they are not what predicates stand for when functioning as predicates. Rather, as we have said, they are the abstract objects denoted by nominalized predicates as abstract singular terms. Predicates, in their functional role as predicates, are not singular terms, and the universals that they stand for in conceptualism are not objects – indeed, they cannot be objects, intensional or otherwise, precisely because of their predicable nature as functional, unsaturated cognitive structures.

There is another notion of property or relation in which properties and relations have a predicable nature, however—but in which, like concepts, they also are not objects. This is the notion in which properties and relations are the unsaturated nexuses or physical structures that underlie the possible states of affairs that can obtain in the world, i.e. the nexuses that are part of the causal structure of the space-time manifold. It is the saturation of properties and relations in this sense that gives material content to the states of affairs that in fact obtain in the world, and it is this material content that is the basis upon which the laws of nature are founded. For it is this notion of a property or relation that is involved in our scientific and theoretical posits about the world—i.e. posits in which such natural properties and relations are assumed to exist (as universals) corresponding to certain of our concepts.

We call the framework in which such posits can be made *conceptual natural realism*, which is the second realist component to what we mean here by conceptual realism as a formal ontology. Thus, with respect to this component, there is, in addition to the primary sense in which a predicate stands for a concept, a secondary sense in which a predicate also stands for the natural property or relation that corresponds to that concept—if, in fact, there is such a corresponding natural property or relation.

We add this last qualification here because not all concepts are assumed to have a natural property or relation corresponding to them. Unlike theoretical concepts, for example, such an assumption does not apply to most of the observational concepts that we employ in our scientific theories. Here,

---

cussion of an alternative reconstruction of Meinong and of how that reconstruction can also be reduced to the system  $\lambda HST^*$ . In Hirst 1989, the importance of having a system of knowledge representation in which the different notions of existence and being that we express in natural language can be expressed is noted and discussed – but Hirst himself seems to have only the Meinongian tradition in mind as a way of achieving this.

the distinction between theoretical concepts and observational concepts is analogous to the historical contrast between primary and secondary qualities. That is, primary qualities, on this analogy, are the concepts that are assumed to have a natural property or relation corresponding to them, and secondary qualities are those that are not. Unlike the historical dichotomy, however, the theoretical-observational distinction between concepts is not absolute and can be applied with more or less rigor, depending on the kind of domain of knowledge representation is in question. In highly theoretical scientific domains, for example, the empirical grounds on which the distinction is made will be generally very rigorous, whereas in the more informal domains of commonsense knowledge, the distinction will have a certain degree of arbitrariness to it, depending on the standards and customs of the domain in question. In either case, what is important is the more general distinction between the order of conception (with its formal categories of thought) and the order of being, both abstract and concrete, as structured by the ontological categories of reality.

This difference between the order of conception and the order of being does not bring with it a difference in logical form between those predicates that stand for concepts and those that stand for natural properties and relations – at least, it does not in conceptual realism *as a formal ontology*. The same predicate can stand in a double way both for a concept (in the primary sense of what determines the truth conditions and correct use of that predicate) and a natural property or relation (in the secondary sense of what, if anything, in nature corresponds to those truth conditions). Similarly, an  $n$ -place predicate variable can be taken in a double way to have both  $n$ -ary concepts and  $n$ -ary natural properties or relations as its values, the difference being determined in context only by whether or not the variable is bound by a quantifier for concepts or a quantifier for natural properties and relations. The latter quantifiers are symbolized here by ‘ $\forall^n$ ’ and ‘ $\exists^n$ ’, which, when affixed to predicate variables, are understood to refer to all, or some, natural properties or relations, respectively. (It is in just this way, but in terms of the quantifiers  $\forall!$  and  $\exists!$ , or the quantifiers  $\forall^e$  and  $\exists^e$ , that we earlier distinguished quantifier phrases that refer only to constructive/predicative concepts, or only to e-concepts, respectively, as opposed to those in terms of  $\forall$  and  $\exists$  that range over concepts in general, predicative or impredicative, or existence-entailing or not.)

The fundamental thesis of natural realism is that every ( $j$ -ary) natural property or relation must, as a matter of natural necessity, be realizable in



nature. This thesis can be stated as follows (with the rule of necessitation for  $\Box^c$  applying to it and other theses of natural realism):

$$(NR) \quad (\forall^n F^j) \Diamond^c (\exists^e x_1) \dots (\exists^e x_j) F(x_1, \dots, x_j).$$

This thesis amounts to a modification of Aristotle's moderate realism, according to which properties and relations can exist only *in re*, i.e. only in the (concrete) objects that have those properties and relations. What  $(NR)$  represents instead is a *modal moderate realism* in which natural properties and relations have a mode of being within the causal structure of the world independently of whether or not there are objects having those properties and relations - but not independently of whether or not there *could* be objects having those properties and relations (in the sense of a causal possibility in nature) (E.g., some properties of transuranic elements will never have instances, because in fact there will never be any atoms of those elements in our world – but, given sufficient technology and energy resources, the possibility of such atoms being caused to exist in accelerators is not precluded by the laws of nature.)

The thesis  $(NR)$  does not apply to concepts, but only to natural properties and relations – a difference that is indicative of the distinction between the conceptual and the ontological orders. (Contradictory concepts, for example, cannot be realized by any existing object.) Another difference between the two orders is that natural properties and relations, but not concepts, are “identical”, if and only if, as a matter of natural necessity, they always have the same instances. This form of cross-world “identity” can be formally defined as follows:

$$(F \equiv_c G) \quad =_{df} \quad \Box^c (\forall x_1) \dots (\forall x_j) [F(x_1, \dots, x_j) \leftrightarrow G(x_1, \dots, x_j)].$$

The assumption that there is a natural property or relation corresponding to a given ( $j$ -ary) concept  $[\lambda x_1 \dots x_j \phi]$ , i.e. that a concept in the one order represents a natural property or relation in the other, can be formulated in terms of  $\equiv_c$  simply as

$$(\exists^n F^j) ([\lambda x_1 \dots x_j \phi] \equiv_c F).$$

This kind of assumption, it must be emphasized, is at best only a scientific hypothesis that can be added as a data structure to particular domains of

knowledge representation, and, unlike instances of the comprehension principle for concepts, it is not to be construed as a logico-conceptual truth.<sup>18</sup>

## 6 Referential Concepts

All of the logical forms introduced so far are perspicuous in the way they specify the truth conditions of formulas in terms of the recursive operations of logical syntax. In this sense logical forms are semantic structures in their own right and do not depend on an external set-theoretic semantics for their interpretation (although such an external semantics may be useful as a guide in proving certain metatheoretic results<sup>19</sup>). This is not all there is to the notion of a logical form in conceptualism, however; for in addition to being logically perspicuous as a semantic structure, a logical form can also be perspicuous conceptually in the way it represents the cognitive structure of a thought or speech act, including, in particular, the referential and predicable concepts underlying that act.<sup>20</sup> Not all logical forms will be perspicuous in this way, it should be noted, because as semantic structures most will be concerned only with the inferential (deductive) part of the overall framework.

Referential and predicable concepts are not the same, but rather are complementary, types of cognitive structures—where the structural complementarity in question is analogous to the way that quantifier phrases are complementary to predicate phrases. That is, predicable concepts, as unsaturated cognitive structures based upon capacities to identify, characterize, and relate objects in various ways, are complemented in speech acts by referential concepts as cognitive structures that enable us to refer (or purport to refer) to the objects that we characterize and relate to one another. And, just as it is the exercise of a predicable concept that informs such an act with a predicable nature, so it is the exercise of a referential concept that informs that act with a referential nature. Similarly, just as it is a predicable concept that

---

<sup>18</sup>See Cocchiarella 1986a, chapter 3, and Cocchiarella 1989a for a detailed development of the logic of natural realism.

<sup>19</sup>See Cocchiarella 1988 for a defense of the distinction between an external and an internal semantics that is based on the more fundamental distinction between a formal theory of predication (such as conceptual realism) and a theory of membership in a set.

<sup>20</sup>This distinction between semantic and cognitive structure, especially in the way the latter is based on referential sortal concepts, is analogous to the distinction in KL-ONE between a formalism for assertion and another for description. See Brachman 1985, p. 174.

underlies our ability to follow the rules of language in the use of a predicate expression, so it is a referential concept that underlies our ability to follow the rules of language in the use of a referential expression. Indeed, strictly speaking, our knowledge of the rules of language for the correct use of different types of expressions is not propositional knowledge but knowledge in the sense of our having these different types of concepts as cognitive capacities; and our following those rules in particular speech acts is just a matter of our exercising those concepts as cognitive capacities. It is this notion of a concept as a rule-following capacity that is also realizable in A.I., moreover—which is why conceptual realism is an ideal framework for knowledge representation in A.I.

Now the complementarity of referential and predicable concepts is syntactically realized in natural language in the difference between noun phrases and verb phrases. Every affirmative assertion (speech act) that is syntactically analyzable in terms of a noun phrase and verb phrase (regardless of the complexity of either) is semantically analyzable in terms of an overt application of a referential and a predicable concept—and the assertion itself is just the result of their mutual saturation in that joint application. A speech act in which ‘All ravens are black’ is asserted, for example, is the result of jointly applying the referential concept that the noun phrase ‘all ravens’ stands for—which, formally, we symbolize as ‘ $(\forall x \text{ Raven})$ ’—with the predicable concept that the verb phrase ‘is black’ stands for—which, formally, is symbolized as ‘ $\text{Black}(\ )$ ’, or as the  $\lambda$ -abstract ‘ $[\lambda x \text{ Black}(x)](\ )$ ’. Thus, in conceptualism, the logical form of the sentence ‘All ravens are black’ is given as ‘ $(\forall x \text{ Raven})\text{Black}(x)$ ’, which represents the cognitive structure of an assertion of that sentence as well as its truth conditions. The logical form of ‘Some ravens are not black’, with the negation understood internally as part of the predicate, is ‘ $(\exists x \text{ Raven})[\lambda x \neg \text{Black}(x)](x)$ ’.

As referential expressions, the objectual quantifier phrases  $(\forall x)$  and  $(\exists x)$  (or their rewrite variants) that we have used so far are really phrases of the form ‘ $(\forall x \text{ Object})$ ’ and ‘ $(\exists x \text{ Object})$ ’, with the category of objects now made explicit as part of the expression. For convenience, we will continue to use the more familiar notation—but only as an abbreviatory device—although, occasionally, we will also use the complete expression to emphasize a point of analysis. In general, accordingly, an objectual quantifier phrase consists not just of a quantifier affixed to an individual variable, but also of a common noun, or what we will call a *common name* (which may be simple or complex).

Grammatically, in English, common names differ from predicates in that

they admit the definite and indefinite articles, plural endings, the adjectives ‘same’, ‘other’, ‘another’, and quantifier words, such as ‘all’, ‘every’, ‘no’, ‘some’, ‘many’, ‘most’, ‘few’, ‘one’, ‘two’, ‘three’, etc. Usually, in a practice going back to Frege, common names are represented in logical syntax as predicates—perhaps because often there are adjectives associated with them in natural language (the way the adjectives ‘canine’, ‘equine’, etc. are associated with the common names ‘dog’, ‘horse’, etc.). We will represent the predicate associated with a common name  $S$  as the  $\lambda$ -abstract  $[\lambda x (\exists y S)(x = y)]$ , which may be read as ‘is (identical with) an  $S$ ’, or, nominalized, as ‘to be (identical with) an  $S$ ’. By  $\lambda$ -conversion, we can abbreviate the application of such a predicate as follows:

$$x S \quad =_{df} \quad (\exists y S)(x = y).$$

The connection between our new logical forms, e.g.  $(\forall x S)F(x)$  and  $(\exists x S)F(x)$ , and the more usual forms is given in the following formulas, which we take as laws of logic (having arbitrary formulas as substituends for  $F$ ):

$$(\forall x S)F(x) \leftrightarrow (\forall x)[x S \rightarrow F(x)],$$

$$(\exists x S)F(x) \leftrightarrow (\exists x)[x S \wedge F(x)].$$

Here it should be noted that despite their logical equivalence, i.e. their equivalence in determining the same truth conditions, the formulas on each side of these biconditionals do not represent the same speech acts—because whereas the speech act represented by the formula on the left side is informed by a referential concept to all (or some)  $S$  and predicates of them that they are  $F$ , the speech act represented by the formula on the right side is informed by a referential concept to all (or some) *objects* and predicates of each of them that *if* it is an  $S$ , *then* it is also  $F$  (or of some object that it is *both* an  $S$  *and*  $F$ ).<sup>21</sup>

Complex common names in English are generated from more basic common names by attaching a relative clause to the latter. We represent the

---

<sup>21</sup>Strictly speaking, in order to represent the cognitive structure of a speech act, the referential concept represented by the quantifier expression occurring in the predicate phrase on the right-hand side of each of these biconditionals must be “deactivated” and intensionalized (through nominalization of the corresponding predicable concept). See Cocchiarella 1989b for a fuller explanation of the difference between an activated and a deactivated referential concept.

operation of attaching a relative clause to a common name by ‘/’, so that if  $S$  is a common name and  $\phi$  is a formula, then ‘ $S/\phi$ ’ is a complex common name (read as ‘ $S$  (who, which) that is (are)  $\phi$ ’). For example, the sentence ‘Every citizen over twenty-one is eligible to vote’ contains the complex common name ‘citizen (who is) over twenty-one’ and can be symbolized (in “longhand”) as

$$(\forall x \text{ Citizen}/x \text{ is over } 21)\text{Eligible} \cdot \text{to} \cdot \text{vote}(x).$$

The logical connection between such complex common names and the more usual logical forms associated with relative clauses is given in the following additional laws of logic (where  $S$  is a schema letter for common names):

$$(\forall x S/\phi)F(x) \leftrightarrow (\forall x S)[\phi \rightarrow F(x)],$$

$$(\exists x S/\phi)F(x) \leftrightarrow (\exists x S)[\phi \wedge F(x)].$$

Iterations of the /-operation can be reduced by means of these laws to simple conjunctions; that is, by the above laws and other standard transformations,

$$(\forall x S/\phi/\psi)F(x) \leftrightarrow (\forall x S/\phi \wedge \psi)F(x),$$

$$(\exists x S/\phi/\psi)F(x) \leftrightarrow (\exists x S/\phi \wedge \psi)F(x),$$

are valid theses of conceptualism.

Finally, it should be noted that singular reference by means of proper names and definite descriptions can be represented as special forms of common name reference. A proper name, for example, can be defined as a common name that necessarily picks out at most one object, and necessarily one and the same object whenever that object exists:

$$PN(S) \quad =_{df} \quad (\forall x S)[\Box^c(\forall y S)(y = x) \wedge \Box^c(E!(x) \rightarrow x S)].$$

(We assume here that  $\Box^c\phi$  implies  $\Box^t\phi$ ). Note that because abstract objects cannot exist, the second conjunct is redundant in the case of names for abstract objects (such as numerals). We should also note that singular reference by means of a proper name can be without, as well as with, existential presuppositions, a distinction that is accommodated in our present notation by the difference in semantic roles between  $(\exists x S)$ , for the use of a proper name  $S$  with existential presupposition, and  $(\forall x S)$ , for the use of  $S$

without such presupposition. A similar distinction can be made for definite descriptions as well.<sup>22</sup>

## 7 The logic of Sortals

Referential concepts are initially developed not in terms of reference to objects in general but to objects of a sort—where, by a *sort* (or *sortal concept*), I mean a type of concept (such as *Raven*, *dog*, *horse*, *car*, *tomato*, etc.) whose use in thought and communication is associated with certain identity criteria, i.e. criteria by which we are able to distinguish and count objects of the sort in question. Sortal concepts are expressed in natural language by common names, both complex and simple.<sup>23</sup>

Not all common names stand for a sort. The common names ‘thing’, ‘individual’, ‘object’ do not, for example, because the concept of a thing, individual, or object *simpliciter* (which is the common name concept to which all others are subordinate), is not associated in its use with any particular identity criteria. In fact, the concept of a thing, individual, or object *simpliciter*, is constructed, according to conceptualism, on the basis of the concept of a sort.<sup>24</sup> Formally, such a construction can be put in the form of a contextual definition according to which a reference to every (or some) object is in effect a reference to every (or some) object of any (or some) sort:

$$(\forall x \text{ Object})\phi \quad =_{df} \quad (\forall S)(\forall x S)\phi,$$

$$(\exists x \text{ Object})\phi \quad =_{df} \quad (\exists S)(\exists x S)\phi,$$

where  $S$  is a variable having sortal concepts (of a given domain) as values. (We will hereafter use  $S$  and  $T$ , with or without subscripts, as sortal variables.)

The above analysis does not commit us to the view that we cannot refer to objects without first being able to identify those objects as being of some specific sort. That is, it allows that we can refer to objects as being of some

---

<sup>22</sup>See Cocchiarella 1989b, section 6.

<sup>23</sup>Sortal concepts are the sortal universals, as opposed to the characterizing universals, of Strawson 1959, and correspond to the substantival, as opposed to the adjectival, general terms of Geach 1962.

<sup>24</sup>Compare Sellars 1963, p. 253f.

sort or other without having any particular sort in mind, or without even being able to specify a particular range of sorts. ('It's a thing of some sort I know not what' is quite meaningful in English.) If it is in fact possible to refer to objects that are in principle not identifiable as being of any sort at all, i.e. if there is a form of reference to objects that the above analysis fails to capture, then it is doubtful that such a form is any way relevant to a framework for knowledge representation in A.I. The above analysis, accordingly, can be taken at least as a methodological thesis, if not also as a conceptual truth of the logic of sortals.

There are many common name concepts other than that of a thing, individual, or object *simpliciter* that are not sortal concepts, i.e. that are not associated with any particular identity criteria. The concept of an artifact, for example, is not a sortal concept – but it is a common-name concept having many different sortal concepts as subordinates. (Consider the problem of trying to write a program for a robot to identify artifacts *simpliciter* on other planets or worlds, i.e. a program that does not specify any particular sort(s) of artifact the robot is to look for.) Similarly, the concept of an event is a common-name concept but not a sortal concept, although, as in the case of artifacts, there are many sorts of events, i.e. there are many subordinate sortal concepts for events.

In general, non-sortal common name concepts to which sortal concepts are subordinate, such as the concept of an artifact, are not definable in terms of a disjunction of their subordinate sortals.<sup>25</sup> Instead, in domains in which such concepts are relevant, they can be represented by primitive common names, and the list of their subordinates can be given in terms of a list of meaning postulates. For example, in a domain in which we want to refer to various sorts of artifacts, we might have some of the following (listing various

---

<sup>25</sup>See Cohen and Murphy 1984, p. 40, where the idea of such disjunctive definitions is also rejected.

sorts of furniture, various means of travel, etc.) as meaning postulates:

$$\begin{aligned}
&(\forall x \textit{ Chair})x \textit{ Artifact}, \\
&(\forall x \textit{ Table})x \textit{ Artifact} \\
&\quad \vdots \\
&(\forall x \textit{ Car})x \textit{ Artifact}, \\
&(\forall x \textit{ Airplane})x \textit{ Artifact}, \\
&\quad \vdots
\end{aligned}$$

Similarly, in a domain in which we want to refer to various sorts of events, we might have

$$\begin{aligned}
&(\forall x \textit{ Kiss})x \textit{ Event}, \\
&(\forall x \textit{ Declaration} \cdot \textit{ of} \cdot \textit{ war})x \textit{ Event}, \\
&(\forall x \textit{ Flash} \cdot \textit{ of} \cdot \textit{ lightning})x \textit{ Event}, \\
&\quad \vdots
\end{aligned}$$

as meaning postulates.

The concept of a physical object (i.e. of a continuant) is also not a sortal, and yet it is important to distinguish physical objects from events, which also exist (as part of the concrete world), i.e. which also are objects that *can* exist. One way to do this is by restricting the meaning of ‘thing’ (as opposed to ‘object’) so as to exclude events, i.e. by assuming the following as a meaning postulate:

$$(\forall x \textit{ Event}) \neg (\exists y \textit{ Thing})(x = y)$$

Abstracts objects are things in this sense, but they can be distinguished from physical things in that only the latter are things that *can* exist (concretely). Reference to physical objects (that do in fact exist), accordingly, can be taken



on this analysis to be reference to *things that exist*. That is, formally,

$$\begin{aligned} (\forall x \text{ Physical} \cdot \text{Object})\phi &=_{df} (\forall x \text{ Thing})(E!(x) \rightarrow \phi), \\ (\exists x \text{ Physical} \cdot \text{Object})\phi &=_{df} (\exists x \text{ Thing})(E!(x) \wedge \phi). \end{aligned}$$

Events are also concrete objects (i.e. objects that *can* exist), but they are different from physical objects in that they are not things.

Aside from their referential role, what is most distinctive about sortal concepts is the role they have in identity statements, particularly in statements of identification and re-identification. A minimal valid thesis about re-identification, for example, is the claim that any object to which we can now refer is identifiable at any other time at which it exists by means of some sortal under which it is now identifiable. Using the now-operator,  $\mathcal{N}$ , this thesis can be formulated as follows:

$$(A) \quad (\forall x)\Box^t (E!(x) \rightarrow (\exists S)[x S \wedge \mathcal{N}(x S)])$$

This thesis is valid even with  $\Box^c$  instead of  $\Box^t$  (and is vacuous when applied to abstract objects).<sup>26</sup> A somewhat stronger thesis, which does not require the now-operator, claims that anything to which we can refer is (now) identifiable by means of a sortal under which it is re-identifiable whenever it exists (even across possible worlds):

$$(B) \quad (\forall x)(\exists S)(x S \wedge \Box^c[E!(x) \rightarrow x S]).$$

Husserl seems to have something like this in mind in his claim that “any object whatever points to a structure ... that is governed by a ... universal rule governing other consciousnesses of it as identical” (1970, p. 53f). If such a rule is to apply only to the object in question, then the claim can be reconstructed in terms of proper names – which are also sortal concepts in that proper names bring with them certain identity criteria associated with their correct use (in a given domain). Stated in terms of proper name sortals, the thesis becomes:

$$(C) \quad (\forall x)(\exists S)[PN(S) \wedge x S],$$

which, given the definition of  $PN(S)$ , implies (B). (Note, however, that whereas (B) is vacuous when applied to abstract objects, (C) is not.) Though

---

<sup>26</sup>See Cocchiarella 1977 for a set-theoretic semantics and more detailed discussion of this and some of the other theses described here regarding the problem of re-identification.

it may be assumed in some special limited domains, we do not propose adopting such a strong thesis for the general framework of knowledge representation. Nor do we propose adopting even the weaker thesis (B)—except, again, at best only for special domains of knowledge representation.

One wrong reason for adopting (C) is the idea that identity across possible worlds (and time) requires a “narrowing” of identity criteria to a point of uniqueness (usually specified in terms of an individual essence). (The notion of an individual essence in these contexts is usually based on a confusion between identification, which is a conceptual process, and individuation, which is an ontological process.) But cross-world and temporal re-identification need not be based on such a “narrowing” of identity criteria, and in fact might be based on a “widening” instead—as provided, for example, by a superordinate sortal. Thus, e.g., it is just such a widening that someone might resort to when he (re-)identifies the girl he met twenty years ago with the woman he has met today: that is, the girl he met years ago is the *same person* as the woman he met today.

Sortal (and common name) subordination (strict and otherwise) can be defined as follows:

$$S_1 \leq S_2 \quad =_{df} \quad \Box^c(\forall x)[x S_1 \rightarrow x S_2],$$

$$S_1 < S_2 \quad =_{df} \quad (S_1 \leq S_2) \wedge \neg(S_2 \leq S_1).$$

One version of the thesis that objects can always be (re-)identified under a superordinate sortal is

$$(D) \quad (\forall S)(\forall x S)\Box^c[E!(x) \rightarrow (\exists T)(S \leq T \wedge x S)],$$

which states that anything identifiable by a sortal  $S$  is identifiable whenever it exists (even across worlds) by some sortal to which  $S$  is subordinate. Note that (D) allows that the superordinate sortal by which an object is identified at one time (or in one world) need not be the same as the superordinate sortal by which that object is identified at a different time (or world)—which indicates why, although (D) implies (A), it does not imply (B). The somewhat stronger thesis that re-identification can be achieved by means of the same superordinate sortal whenever the object exists, i.e.,

$$(E) \quad (\forall S)(\exists T)(S \leq T \wedge (\forall x S)\Box^c[E!(x) \rightarrow x T]),$$

does imply (B) (and therefore (A) as well). But as with (B), we do not claim here that either (D) or (E) is valid without qualification in all domains of

knowledge representation. Such qualification can be found in the framework of conceptual natural realism in the ontological category of natural kinds.

## 8 Sortals and Natural Kinds

In natural language and our commonsense understanding of the world there is a basic dichotomy between sortals for artifacts and sortals for natural kinds – where, by a natural kind, I mean not a concept but a causal structure of a certain determinate type or category. In particular, a natural kind is a type of causal structure or mechanism in nature that is the basis of the powers or capacities that objects have to act, behave, function, etc., in certain determinate lawlike ways—i.e. ways that lead us to speak of those objects as belonging to a natural kind. It is the ontological category of natural kinds as causal structures that is the basis of the natural laws regarding the different natural kinds of objects there are, or could be, in the world.

Informally, natural kinds are understood here to include the various genera and species of plants and animals—which we will speak of as natural kinds of *things* – as well as the natural kinds of substances (in the chemical sense) such as gold, water, oxygen, etc., which we will speak of as natural kinds of *stuff*. Natural kinds of things (in the sense understood here) are wholes that are more than the sum of their parts or the bits of stuff that make them up (and correspond to what Aristotle meant by primary substances). An animal or plant, for example, is a natural kind of thing that over time will change much of the material stuff that makes it up, while still remaining the same animal or plant. A puddle of water, on the other hand, or the sodium chloride in a salt shaker, etc., is also a (concrete) thing (and as such is a value of the bound individual variables), but it is so only in the sense of being an actual bit (or quantity) of a natural kind of stuff, and as such it is no more than the sum of the minimal bits of the kind(s) of stuff that make it up. (The minimal bits of a natural kind of stuff are atoms or molecules, which are the only bits of that kind that are natural kinds of things as well. That is, atoms and molecules of a natural kind of stuff are both minimal bits of that kind of stuff and natural kinds of things.) As already noted, natural kinds of stuff include compound substances, such as water or salt, as well as elementary substances, such as hydrogen, gold, iron, uranium, etc.

The assumption that there are natural kinds of things or stuff is a development of conceptual natural realism into a version of Aristotelian essentialism

– where natural kinds correspond to what Aristotle called secondary substances. Aristotelian essentialism, or something very much like it, is implicit in the commonsense view of the world that we express in natural language—a framework in which natural kinds, as already noted, are taken to correspond not to predicates but to sortal common names. Ontologically, this means that the category of natural kinds is not the same as the category of natural properties—a categorial distinction that we retain in conceptual realism as a formal ontology, where natural kinds are assumed to correspond not to predicable concepts (as I had maintained in earlier articles) but to sortal concepts instead. This approach will provide not only a more appropriate representation of the ontology of natural language understanding, but also a clearer metaphysical foundation for our theoretical sciences as well. For this reason, accordingly, we introduce new quantifiers  $\forall^k$  and  $\exists^k$ , which, when affixed to sortal variables, are assumed to refer to natural kinds.

The question as to which of our sortal concepts there corresponds a natural kind is no less an empirical matter than is the question as to which of our predicable concepts there corresponds a natural property or relation, and the grounds on which we make such an assumption can be either very rigorous (as in highly theoretical domains) or very informal (as in the domains of commonsense knowledge or belief). Also, just as in the case for natural properties and relations, there are some general laws for natural kinds that apply to all domains—i.e. laws of conceptual realism that are part of the general framework for knowledge representation. Thus, for example, just as it is part of natural realism in general that every natural property or relation is realizable in nature, so too every natural kind is realizable in nature as well (even though natural kinds are not properties but causal structures complementary to properties), i.e.,

$$(K1) \quad (\forall^k S) \Diamond^c (\exists^e x) x S,$$

is a general law for natural kinds. Another such general law is that only (concrete) existents can belong to a natural kind, i.e.,

$$(K2) \quad (\forall^k S) (\forall x) [x S \rightarrow E!(x)],$$

in which case the restricted existential quantifier  $(\exists^e x)$  in  $(K1)$  can be replaced by the unrestricted existential quantifier  $(\exists x)$ . Also, just as natural properties are “identical” when, as matter of natural necessity, they always have the same instances, so too natural kinds are “identical” if, as a matter

of natural necessity, any object belonging to the one kind must belong to the other as well. We can, accordingly, extend our use of  $\equiv_c$  as an abbreviatory notation so as to apply to sortal expressions as well:

$$(S_1 \equiv_c S_2) \quad =_{df} \quad (S_1 \leq S_2) \ \& \ (S_2 \leq S_1).$$

The thesis, which also holds for natural properties and relations, that natural kinds are the same in all causally possible worlds (i.e. worlds having the same laws of nature as ours) is expressed with this notation as follows:

$$(K3) \quad (\forall^k S_1) \Box^c (\exists^k S_2) (S_1 \equiv_c S_2).$$

The analysis of what it means to say that there is a natural kind corresponding to a given sortal  $S$  (complex or simple), i.e. that  $S$  represents a natural kind, can be given now simply as follows:

$$NK(S) \quad =_{df} \quad (\exists^k T) (S \equiv_c T).$$

There are general laws for natural kinds that do not also hold for natural properties. The most important and fundamental such law is a law of rigidity according to which an object can belong to a natural kind only if it *must* belong to that kind whenever it exists, i.e.,

$$(K4) \quad (\forall^k S) (\forall x) (\Diamond^c (x \ S) \rightarrow \Box^c [E!(x) \rightarrow x \ S])$$

(With S5 as the modal logic for  $\Box^c$  and  $\Diamond^c$ , the antecedent condition,  $\Diamond^c (x \ S)$ , can be replaced by just  $(x \ S)$ .) Note that, by (K4), any object that is identifiable by means of a sortal to which there corresponds a natural kind is re-identifiable by means of that sortal whenever it exists. That is,

$$(B^k) \quad NK(S) \rightarrow (\forall x \ S) \Box^c [E!(x) \rightarrow x \ S],$$

which is the counterpart of (B) for objects identifiable by means of a sortal for a natural kind, is provable on the basis of (K4). If we also assume (as an ideal of scientific inquiry) the methodological thesis that every object belonging to a natural kind is identifiable by means of a sortal that represents a natural kind, i.e.,

$$(MT1) \quad (\forall x) [(\exists^k T) x \ T \rightarrow (\exists S) (NK(S) \ \& \ x \ S)],$$

then an identifying principle even closer to  $(B)$ , namely,

$$(B_2^k) \quad (\forall x)[(\exists^k T)(x T) \rightarrow (\exists S)(x S \ \& \ \Box^c[E!(x) \rightarrow x S])],$$

is also derivable from  $(K4)$ .

There are other principles of (re-)identification that, like  $(B)$ , are not valid in general, but which are valid when suitably restricted to sortals for natural kinds. One such principle is the intersection thesis of Stevenson 1975 that if any two sortals intersect then there is a sortal to which both are subordinate:

$$(IT) \quad (\exists x S_1)x S_2 \vee (\exists x S_2)x S_1 \rightarrow (\exists T)(S_1 \leq T \wedge S_2 \leq T).$$

This thesis is not valid in general because new sortals can always be added to a given domain of knowledge representation that can falsify it, so that even if it were to hold in a given restricted domain, it need not also hold in an expanded or larger domain. With an ontology of natural kinds, however, i.e. with constraints that go beyond mere concept-formation, an even stronger version of this thesis is valid for sortals suitably qualified.

The relevant ontological law here is the *Partition Principle*, i.e., the principle that if two natural kinds that can intersect, then one must be subordinate to the other:

$$(K5) \quad (\forall^k S_1)(\forall^k S_2)[\Diamond^c(\exists x)(x S_1 \wedge x S_2) \rightarrow S_1 \leq S_2 \vee S_2 \leq S_1].$$

The idea behind this principle is that the family of natural kinds to which an object can belong forms a chain of subordination of one natural kind to another - where each natural kind in the chain is, as it were, a template structure that is causally more determinate and finer-grained than the natural kinds to which it is subordinate. The partition determined by this chain is described in a consequence of  $(K5)$ , which states in effect that natural kinds that are subordinate to the same immediate genus are either “identical” or necessarily disjoint.<sup>27</sup> Together with  $(K4)$ , incidentally,  $(K5)$  implies that an object can belong to two natural kinds only if, as a matter of natural necessity, it belongs to the one kind when and only when it belongs to the other, i.e.,

$$(\forall^k S_1)(\forall^k S_2)(\forall x)[\Diamond^c(x S_1) \wedge \Diamond^c(x S_2) \rightarrow \Box^c(x S_1 \leftrightarrow x S_2)]$$

---

<sup>27</sup>See Cocchiarella 1989a, section 14, for a detailed formulation of this and other laws for natural kinds.

is a valid consequence of (K4) and (K5). In regard to the intersection thesis, note that if  $S$  and  $T$  are sortals for which there are corresponding natural kinds, then, by (K5), the following strong intersection thesis,

$$NK(S) \ \& \ NK(T) \rightarrow [\Diamond^c(\exists x \ S)x \ T \vee \Diamond^c(\exists x \ T)x \ S \rightarrow S \leq T \vee T \leq S].$$

is valid.

Another assumption made in Stevenson 1975 that is not universally valid for sortals, but which has some plausibility when restricted to sortals for natural kinds, is the thesis that every sortal concept is subordinate to some ultimate sortal. Defining the notion of an ultimate sortal as follows,

$$Ult(S) \quad =_{df} \quad \neg(\exists T)(S < T),$$

Stevenson's ultimate-sortals thesis can be formulated as

$$(US) \qquad (\forall S)(\exists T)[S \leq T \wedge Ult(T)].$$

Like (IT), this thesis is not valid for all domains, because, unlike the kinds of constraints imposed by ontological laws, there are no laws or constraints on concept-formation regarding the addition of new sortals to a given domain of knowledge representation. That is, new sortals can always be added to a given domain that would invalidate (US) in an extended or larger domain.

What gives (US) its plausibility is the qualified use it can have for sortals that represent natural kinds. A sortal for an ultimate natural kind, for example, can be defined as follows,

$$Ult^k(S) \quad =_{df} \quad NK(S) \ \& \ \neg(\exists T)[NK(T) \wedge S < T],$$

and a summum genus principle to the effect that every natural kind sortal is subordinate to an ultimate natural kind sortal can be formulated as

$$(US^k) \qquad (\forall S)(NK(S) \rightarrow (\exists T)[Ult^k(T) \wedge S \leq T]).$$

Note that a restricted version of thesis (E) (and of (D) as well) of the previous section, namely,

$$(E^k) \qquad (\forall S)[NK(S) \rightarrow (\exists T)(S \leq T \wedge (\forall x \ S)\Box^c[E!(x) \rightarrow x \ T])],$$

is a consequence of (US<sup>k</sup>).

This does not mean that  $(US^k)$  and  $(E^k)$  are valid. What must be distinguished here is  $(US^k)$  as a conceptual thesis about sortals for natural kinds, as opposed to an ontological thesis about natural kinds *simpliciter*, i.e. about natural kinds independently of whether or not they are represented by sortals. The relevant ontological law here is not  $(US^k)$ , but the *Summum Genus Principle* that every object belonging to a natural kind belongs to a summum genus of all the natural kinds to which that object belongs, i.e.,

$$(K6) \quad (\forall x)(\forall^k S_1) (x S_1 \rightarrow (\exists^k T)[x T \wedge (\forall^k S_2)(x S_2 \rightarrow S_2 \leq T)]) .$$

Given the partition principle  $(K5)$ , an alternative way to formulate  $(K6)$  is as

$$(K6') \quad (\forall^k S_1)(\exists^k T)[S_1 \leq T \wedge \neg(\exists^k S_2)(T < S_2)],$$

which is the ontological counterpart of  $(US^k)$  as a conceptual thesis.

One way to connect the ontological law and the conceptual principle here is by means of a methodological assumption that (as an ideal of scientific inquiry) every ontological summum genus, i.e. every natural kind that is a summum genus, can be represented by a sortal:

$$(MT2) \quad (\forall^k S_1)[\neg(\exists^k S_2)(S_1 < S_2) \rightarrow (\exists T)(S \equiv_c T)].$$

That is, given  $(K6')$  as an ontological law of conceptual realism, and  $(MT2)$  as a methodological thesis,  $(US^k)$  as a conceptual thesis follows.

There are other ontological laws for natural kinds – such as an *Infima Species Principle*, or the principle that every genus is the sum of its species, etc. – that are also part of the general framework of conceptual realism as a formal ontology. We forego their discussion here and note only that these laws also have conceptual counterparts, which while not valid in general, may have qualified forms in which they may have at least a certain methodological utility. It is this distinction between ontological laws and conceptual theses, and the methodological issues that connect them, that seems to be the real basis of the distinction between what is called the ontological level and the conceptual level – and perhaps also the epistemological level as well—in knowledge representation.

## 9 Concluding Remarks

We have given here only an overview or sketch of conceptual realism as a formal ontology, i.e. as a theory of logical form having both conceptual and



ontological categories—but in which the latter are represented in terms of the former. The categories of natural kinds and of natural properties and relations, for example, are represented in terms of the categories of sortals and predicable concepts, respectively, and the category of abstract objects is represented in terms of the process of conceptual nominalization (reification) as a subcategory of objects. Not all of these categories or parts of this formal ontology will be relevant in every domain of knowledge representation, but each is relevant at least to some domains and is needed in a comprehensive framework for knowledge representation. In those domains where certain categorial distinctions are not needed—such as that between predicative and impredicative concepts, or that between predicable concepts and natural properties and relations, or between sortal concepts and natural kinds, etc.—we can simply ignore or delete the logical forms in question. What must remain as the core of the system is the intensional logic around which all of the other categories are built—namely, the second-order predicate logic with nominalized predicates as abstract singular terms that we call  $HST_{\lambda}^*$ . It is this core, I believe, that can serve as a universal standard by which to evaluate other representational systems.

## REFERENCES

- Bolzano, B. 1972 *Theory of Science*, translated by R. George (Berkeley: University of California Press).
- Brachman, R.J. 1985 “An Overview of the KL-ONE Knowledge Representation system”, *Cognitive Science* 9: 171-216.
- Chierchia, G. 1984 *Topics in the Syntax and Semantics of Infinitives and Gerunds*, Ph.D. dissertation, University of Massachusetts.
- Chierchia, G. 1985 “Formal Semantics and the Grammar of Predication”, *Linguistic Inquiry* 16: 417-443.
- Church, A. 1956 *Introduction to Mathematical Logic*, vol. 1 (Princeton: Princeton University Press).
- Cohen, B. and Murphy, G.L. 1984 “Models of Concepts”, *Cognitive Science* 8: 27-58.

- Cocchiarella, N.B. 1977 “Sortals, Natural Kinds, and Re-Identification”, *Logique et Analyse* 80: 339-474.
- Cocchiarella, N.B. 1984 “Philosophical Perspectives on Quantification in Tense and Modal Logic”, in *Handbook of Philosophical Logic*, vol. 2, D. Gabbay and F. Guenther, eds. (Dordrecht: D. Reidel): 309-353.
- Cocchiarella, N.B. 1986a *Philosophical Investigations of Predication Theory and the Problem of Universals* (Naples: Bibliopolis).
- Cocchiarella, N.B. 1986b “Conceptualism, Ramified Logic, and Nominalized Predicates”, *Topoi* 5: 75-87.
- Cocchiarella, N.B. 1987 *Logical Studies in Early Analytic Philosophy* (Columbus: Ohio State University Press).
- Cocchiarella, N.B. 1988 “Predication versus Membership in the Distinction Between Logic as Language and Logic as Calculus”, *Synthese* 77: 37-72.
- Cocchiarella, N.B. 1989a “Philosophical Perspectives on Formal Theories of Predication”, in *Handbook of Philosophical Logic*, vol. 4, D. Gabbay and F. Guenther, eds. (Dordrecht: D. Reidel): 253-326.
- Cocchiarella, N.B. 1989b “Conceptualism, Realism, and Intensional Logic”, *Topoi* 8: 15-34.
- Cocchiarella, N.B. 1991 “Quantification, Time, and Necessity”, in *Philosophical Applications of Free Logic*, K. Lambert, ed. (Oxford: Oxford University Press): 242-256.
- Geach, P. 1962 *Reference and Generality* (Ithaca: Cornell University Press).
- Hirst, G. 1989 “Ontological Assumptions in Knowledge Representation”, in *Principles of Knowledge Representation and Reasoning*, R. Brachman, H.J. Levesque, and R. Reiter, eds. (San Mateo: Morgan Kaufmann Pub.): 157-169.
- Hobbs, J.R. and Moore, R.C. 1985 *Formal Theories of the Commonsense World* (Norwood, N.J.: Ablex).
- Husserl, E. 1970 *Cartesian Meditations*, translated by Dorian Cairns (The Hague: Martinus Nijhoff).

Landini, G. 1990 “How to Russell Another Meinongian: An Early Russellian Account of Fictional Objects versus Zalta’s Abstract Objects”, *Grazer Philosophische Studien* 37: 93-122.

Montague, R.M. 1974 *Formal Philosophy* (New Haven: Yale University Press).

Parsons, T. 1980 *Nonexistent Objects* (New Haven: Yale University Press).

Prior, A. 1968 *Past, Present, and Future* (Oxford: Oxford University Press).

Sellars, W. 1963 “Grammar and Existence: A Preface to Ontology”, in *Science, Perception, and Reality* (London: Routledge & Kegan Paul).

Shapiro, S. 1985 *Intensional Mathematics* (Amsterdam: No. Holland Pub. Co.).

Strawson, P. 1959 *Individuals, An Essay in Descriptive Metaphysics* (London: Methuen & Co.).